

# Using eye-tracking to study anaphoric relations processing in European Portuguese

<sup>1</sup>A Costa, G Matos, P Luegi

Faculdade de Letras da Universidade de Lisboa – CLUL, PORTUGAL

<sup>1</sup>armandacosta@fl.ul.pt

## ABSTRACT

In this study we evaluate processing costs of different types of anaphoric expressions during reading. We consider three types of anaphoric expressions in Subject sentential position: a null pronoun (*pro*), and two gaps produced by syntactic movement: a WH-variable and a NP copy. Given that coreferential *pro* exhibits more referential weight than wh- and NP-gaps, and grounded on theories of referential processing based on relations of hierarchy and accessibility of the antecedent, we raise the hypothesis that the more dependent on its antecedent the anaphoric null constituent is, and the more minimal is the distance in terms of hierarchical structure between the anaphoric null element and its antecedent, the lower are the cognitive costs in processing. To test our hypothesis, we registered the eye movements with R6-HS ASL system of 20 Portuguese adult native speakers. Text regions including the selected anaphoric expressions were delimited and tagged. We analyzed the reading time of each region taking into account the number and duration of eye fixations per region; we used the reading time by character in milliseconds in order to compare values between regions of different length. We found a significant advantage in the reading time of the gaps arising from movement over the reading time of *pro*.

## 1. INTRODUCTION

When listening to a discourse or reading a text, the listener or reader has to process a lot of information about discursive entities that appear in a discontinuous way. This is particularly evident in narratives when there is a set of characters that, after the first mention, are referred to throughout the story. For the sake of discourse cohesion, it is necessary to take anaphoric expressions to refer to the characters by means of appropriate anaphoric expressions, avoiding redundancy and ambiguity. If more than one character participates in a given story, it will be necessary to refer to them as the story unfolds, in such a way that it will be clear for the listener or reader what is the character under reference. For instance, in the *Cat Story*, used crosslinguistically to study the reference mechanisms to person, space and time (Hickmann, 1995; Batoréo, 1996, among others), we have three entities - a cat, a dog and a bird - that participate in the events' chain. When the speaker or writer has to refer to one of them, he must choose the most appropriate linguistic element in order to cue the listener or reader in the good way to understanding. Studying the production of oral and written narratives in European Portuguese (EP), we found out that ten years old children were already able to use linguistic mechanisms to express reference in a cohesive way (Batoréo and Costa, 1999). We verified that, at this developmental stage, children make use of all the linguistic possibilities to create well formed referential chains. However, there is a specific use of certain expressions depending on the oral or written modality of the language. In particular, in oral narratives there is a preference for more economic forms with deictic content, such as pronouns - *ele/it* or *zero anaphora* (*pro*) - while in writing there is a preference for noun repetition or semantically related nouns. This shows that the language modality has an impact on the referential mechanisms adopted.

Languages share linguistic processes to mark coreferential relations in discourse. We can retrieve a discursive entity through lexical or pronominal means. In the first case, a first mentioned entity can be referred by iteration - for example: *the cat was there, the cat tried to catch the bird, the cat run away* - or by a semantic related noun: a hyperonym - *the animal, the feline* - or by a hyponym - *a pet name* for example; the reference can also be made by a modified or an enriched noun phrase - *the cat that was trying to catch the bird*. In a more economical way, pronouns can be employed. In this case, the pronominal form may signal more or less unequivocally its relationship with the antecedent depending

on the shared morphological information that can provide clues of gender or number – *he/she/it* or *they*; in the case of demonstrative (DEM) pronoun, *this/that*, it points out the spatial and deictic relation with the antecedent.

As we have seen, the inventory of anaphoric forms that can be used is extended. In this study we will analyze a subset of the available possibilities in European Portuguese to create anaphoric chains, involving the most economical anaphoric elements: null (or covert) pronouns and gaps arising from syntactic movement.

## 2. ANAPHORIC RELATIONS PROCESSING IN EUROPEAN PORTUGUESE

### 2.1 Referential processing in European Portuguese

Theories of hierarchy and accessibility to explain reference assignment (*The accessibility marking scale*, Ariel 1996, 2001; *Centering theory*, Grosz, Joshi and Weinstein, 1995) predict that processing strategies are held up by the phonetic, morphological and syntactic form of the anaphoric expression, which works as a cue to search the antecedent. In an anaphoric chain, its head, that is, the first mentioned entity, could have different levels of activation in working memory due to its structural position in the sentence, its thematic role, or its prominent pragmatic status. When the antecedent is simultaneously a Subject, an Agent and a Topic (thus occupying the sentence first position), as in (1) and (2), it displays a set of conditions that turn it into the best identifier of an anaphoric category, occurring the sentence subject position in a more embedded sentence (cf. (1)) or in the next discursive fragment (Costa, Faria and Kail, 2004; Costa, 2005, 2010; Costa, Matos and Faria, 1998; Costa, Matos and Luegi, 2010; Luegi, in progress; Morgado in progress).

- (1) O João<sub>i</sub> cumprimentou o Pedro amistosamente quando [pro]<sub>i/?j</sub> entrou no restaurante.  
*John greeted Peter warmly when [he] went into the restaurant.*
- (2) O João<sub>i</sub> cumprimentou o Pedro amistosamente. Ele<sub>i/?j</sub> estava no restaurante.  
*John greeted Peter warmly. He was in the restaurant.*

Grammar and processing studies on EP show that, in Null Subject languages like Portuguese, Italian or Spanish, when co-reference is at stake, an antecedent with a syntactic or discursive prominence favours the occurrence of a null pronominal in subject position, or, in more technical terms, a *pro* (Brito, 1991; Carminati 2002; Alonso-Ovale et al, 2002). In complex sentence, this preference takes place over an overt pronominal inside the embedded domain. This has been interpreted as a consequence of an economy strategy recalling the *Avoid Pronoun Principle* originally proposed in Chomsky, 1981. This is also in accordance with the preference for minimal forms to point out prominent antecedents (Ariel, 1996, 2001; Grosz et al, 1995). Sentences (3) exemplify these conditions.

- (3a) O João<sub>i</sub> não cumprimentou o Pedro porque [pro]<sub>i</sub> estava com a sua ex-namorada.  
*John didn't greet Peter because [he] was with his ex-girlfriend.*
- (3b) O João<sub>i</sub> não cumprimentou o Pedro<sub>j</sub> porque ele<sub>i/?j</sub> estava com a sua ex-namorada.  
*John didn't greet Peter because he was with his ex-girlfriend.*

### 2.1 European Portuguese grammatical properties and null anaphoric expressions

European Portuguese exhibits interesting conditions for the study of the processing of anaphoric relations since it is a null subject language. In this language *pro* can occur as the Subject of a simple clause designating a definite entity, as in (4), and also as the Subject of an embedded clause in a complex sentence, participating in anaphoric relation with a previous NP, as in (5).

- (4) [pro]<sub>1st person; plural</sub> Lemos esse livro recentemente.  
*(We) read that book recently.*
- (5) [O homem]<sub>i</sub> levantou-se do degrau quando [pro]<sub>i</sub> ouviu correr os ferrolhos.  
*The man stood up from the step when (he) heard to open the bolts.*

Some grammatical properties characterize this null category in subject position in Null Subject Romance languages: *pro* is a pronoun basically inserted in an argument position, locally licensed by the verbal inflexional agreement, which may identify its content, whenever 1st and 2nd person agreement features are available (cf. (4)), otherwise it is the linguistic context (cf. (3) and (5)) or the pragmatic situation that sets its referential value. If *pro* depends on the linguistic context to retrieve its value, it participates in a referential chain, whose content is established by its antecedent (5). When it enters in a co-reference relationship, it becomes a member of a complex referential chain involving two independent referential elements (the antecedent Noun Phrase and *pro*).

However, in Subject position, in EP as in other languages, we can find other types of syntactic objects without phonological form that enter into an anaphoric relationship. In (6a), we have an embedded relative clause, which has a gap as its Subject (a property that becomes evident when another clause intervenes between the relativized constituent and the clause containing the gap, as in (6b)). In (6a), the gap must be interpreted in relation to its antecedent that is a WH- expression (*que* ‘who/that’) that binds it.

- (6a) O homem<sub>i</sub> [que<sub>i</sub> [-]<sub>i</sub> queria um barco] foi bater à porta do rei.  
*The man who wanted a boat went to the king's house.*
- (6b) O homem<sub>i</sub> [que<sub>i</sub> todos diziam que [-]<sub>i</sub> queria um barco] foi bater à porta do rei.  
*The man who everybody said that wanted a boat went to the king's house.*

In this case we have a null category that, unlike *pro*, arises from the displacement of a WH-phrase to a left peripheral position of the sentence, that is, to a non-argument position. This gap is a variable bound by the displaced WH-phrase and has not independent referential content.

In (7), we have sentence coordination. We assume that, in consequence of an *across-the-board movement* of the NP-phrase in subject position of each coordinate sentence raises into a single position external to the coordinate structure, leaving a NP-gap inside each conjunct. These NP-gaps occur in argument position and their content is established by the displaced NP, which locally binds them (Matos, 2001; Colaço, 2005).

- (7) O homem<sub>i</sub> [[-]<sub>i</sub> levantou-se, [-]<sub>i</sub> enrolou a manta e [-]<sub>i</sub> pôs-se à espera].  
*The man stood up, wrapped the blanket and waited.*

Thus, we have three types of null categories that can occur in Subject position entering in an anaphoric relation: the covert pronominal (*pro*), the WH-variable and the NP-gap. This opens the range of the syntactic nominal null categories that may occur in the Subject position of a finite clause in EP and constitutes a challenge to processing referential theories which do not predict differences in processing those types of syntactic categories.

### 3. EXPERIMENTAL STUDY

#### 3.1 Hypothesis

Considering the economy requirements of the computational system of the grammar and the processing costs involved in a comprehension task, we raise the hypothesis that the less referential an anaphoric null category is, and the more minimal the distance or structural complexity between the antecedent and the null element are, the lower are the cognitive costs. Thus, considering *pro*, WH- and NP-gaps, we predict that processing costs will be higher, in terms of reading time, in contexts where the *pro* gaps occur.

For cognitive effort we understand the overload of the working memory due to difficulties in integration processes. The maintenance in memory of the members of a referential chain must have an effect on the load of the working memory depending on the distance or the relationship between the elements. It is expected that this effect will be reflected in some behavioural measures that can be registered with an eyetracker equipment.

#### 3.2 Procedure

The study of anaphoric relation's processing implies to get behavioral indicators of the cognitive and linguistic processes triggered by the interpretation of a linguistic chain formed by a nominal head and one or more anaphoric expressions that refer to it in a backward (occasionally forward) relation. In

psycholinguistic research the main objective is to capture the mental operations that take place in the online processing and to relate them with the grammatical properties of the verbal input.

Eye tracking has been intensively used in the last decades to study the online sentence processing (Rayner, 1998). This is based on the assumption that there is a correspondence between the cognitive processes that occur during reading and the eye movements which can be recorded by the eyetracker and converted into dependent variables, such as fixations' duration, gaze duration, first pass, second pass, and so on (Just and Carpenter 1980; Hyona, Lorch Jr and Rinck 2003; Boland 2004; Rayner and Pollatsek, 2006; Rayner, Juhasz and Pollatsek 2007). There is the conviction that this methodological procedure provides an online measure of the mental operations during processing with high temporal resolution. Against other methodologies (self-paced reading, for instances), eyetracking does not rely on secondary tasks that could interfere in the main aim: reading and comprehension. To study reading, eye fixations and saccadic movements are the relevant behaviours for analysis.

In the present study we measured fixations' location, order and duration. We had to analyse how the readers deal with complex anaphoric chains that are located in a text in a non-predictable way; we want to verify how the different syntactic status of the anaphoric elements affects processing costs. In order to accomplish this goal, we used an eye tracking paradigm, taking as dependent measure the time spent in reading each region where the target anaphoric element occurs.

A sample of 20 Portuguese adult native speakers, undergraduate students of the University of Lisbon, classified as good readers by means of a questionnaire on reading habits, participated in the experiment. The eye-movements of the participants were registered with the R6-HS model of ASL (pupil and corneal reflection, with an accuracy of 0.5 degree of visual angle, data collected at a rate of 240 Hz) while they read an excerpt of the novel *The Tale of the Unknown Island*, written by José Saramago. The text, which had 809 words, was presented on a screen divided in 9 different slides. After calibrating the equipment, the participants had to read an introductory slide, with biographical information about the writer, in order to adapt them to the task and also to verify if the equipment was well calibrated. Once the reading task completed, the participants had to fulfill a questionnaire, asking for information about the text. The purpose of this task was to assure that the participants read the text with attention intending to understand it (subjects were warned of the existence of this task from the beginning).

For analysis purpose, we selected some fragments in the text, which included complete anaphoric chains. These fragments were analyzed syntactically and tagged properly. In order to proceed with the analysis of the register of eyetracking, each fragment was divided in several regions, each one containing an anaphoric element at the left border of the clause.

In Table 1 we present the five selected fragments, and, within them, the delimited regions that incorporate the target anaphoric expression (the vertical bar marks the borders of the region). The left

*Table 1 – text fragments analyzed in terms of eye fixations in each marked region*

---

(8)	Este chamava o terceiro <sub>i</sub> ,   [CP que <sub>i</sub> [TP [WH-] <sub>j</sub> mandava o primeiro-ajudante <sub>j</sub> ,   [CP que <sub>j</sub> [por sua vez] [TP [WH -] <sub>j</sub> mandava o segundo]]]   <i>this called the third one, who ordered the first-helper, who in turn ordered the second (one)</i>
(9)	até [pro] chegar à mulher da limpeza <sub>j</sub> ,   [a qual <sub>j</sub> [não [pro] <sub>j</sub> tendo ninguém   [CP em quem [[pro] <sub>j</sub> mandar]]   [TP [WH -] <sub>j</sub> entreabria a porta e [[WH -] <sub>j</sub> perguntava]]. <i>until the cleaning lady, who, not having anybody (who) to order opened the door and asked</i>
(10)	ao primeiro-secretário <sub>i</sub> ,   [o qual <sub>i</sub> , [[pro] escusado seria dizer],   [IP [WH -] <sub>i</sub> passava a encomenda ao segundo-secretário]]   <i>to the first-helper who, needless to say sent the parcel to the second-helper,</i>
(11)	até [pro] <sub>j</sub> chegar à mulher da limpeza <sub>j</sub> ,   [CP que <sub>j</sub> [TP [WH-] <sub>j</sub> despachava sim ou não   conforme [pro] <sub>j</sub> estivesse de maré ] <i>until the cleaning lady, who replied yes or no according (to) her mood</i>
(12)	TP[O homem   [CP que <sub>i</sub> [TP [WH-] <sub>i</sub> queria um barco]]] [TP1[NP-] <sub>i</sub> levantou-se do degrau da porta

---

| [quando [pro]<sub>i</sub> começou a ouvir correr os ferrolhos]], | [TP2 [NP-]<sub>i</sub> enrolou a manta]

| e [TP3 [NP-]<sub>i</sub> pôs à espera]].

*the man who wanted a boat got up from the step of the door when (he) started to hear (to) open the bolts, rolled up his cloak and waited*

CP=complementizer phrase; TP=Tense Phrase; [pro]=null pronominal; [WH-]=WH-gap; [NP-]=NP-gap

border of each region matches the boundary of a clause, which begins with a connector followed by an anaphoric expression. This anaphoric expression is both Subject and (non-marked) Topic of the clause and co-refers with an antecedent presented in a previous clause.

In order to compare processing costs of each type on anaphoric element, in each region, we measure:

- The Number of Fixations (NFix) – number of all fixations done in each region;
- The Mean Fixation Duration (MFD) – mean of the duration of all fixations done in a region;
- The Total Reading Time (TRT) – corresponding to the sum of the durations of all fixations done in a region (both in first and second-pass);
- The Total Reading Time per Character (TRT/Char) – total reading time (seconds) spent on a region divided by its number of characters and white spaces.

### 3.3 Results

In Table 2, we present the results for each fragment and each region. This study is not a normative one, since the linguistic material is authentic text. As a consequence, the regions under analysis vary a lot in length. That's why we used mainly the Reading Time per Character for explain the results and discuss our hypothesis.

**Table 2 – results for each fragment and each region**

(1)	o primeiro- secretário chamava o segundo- secretário	este chamava o terceiro	que mandava o primeiro-ajudante	que por sua vez mandava o segundo,
<i>N.Car</i>	NP <sub>i</sub>	DEM <sub>i</sub>	WH <sub>j</sub>	WH <sub>k</sub>
	51	24	31	34
TRT/Char	<b>0.041</b>	<b>0.057</b>	<b>0.046</b>	<b>0.053</b>
	0.013	0.018	0.015	0.021
MFD	<b>0.258</b>	<b>0.276</b>	<b>0.264</b>	<b>0.253</b>
	0.054	0.059	0.046	0.054
NFix	<b>8.3</b>	<b>5.1</b>	<b>5.4</b>	<b>7.0</b>
	2.7	1.7	1.6	1.9

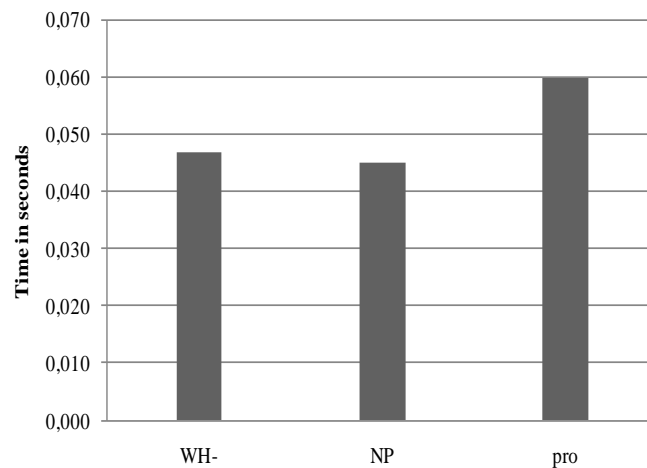
(2)	até chegar à mulher da limpeza,	a qual, não tendo ninguém	em quem mandar,	entrebria a porta	e perguntava pela frincha,
<i>N.Car</i>	NP <sub>i</sub>	WH <sub>i</sub>	WH <sub>j</sub>	pro <sub>i</sub> V	pro <sub>i</sub>
	31	26	15	19	26
TRT/Char	<b>0.043</b>	<b>0.061</b>	<b>0.058</b>	<b>0.055</b>	<b>0.054</b>
	0.015	0.030	0.029	0.027	0.019
MFD	<b>0.275</b>	<b>0.266</b>	<b>0.246</b>	<b>0.355</b>	<b>0.254</b>
	0.075	0.059	0.055	0.139	0.054
NFix	<b>5.0</b>	<b>6.0</b>	<b>3.5</b>	<b>3.2</b>	<b>5.6</b>
	1.9	2.6	1.6	1.7	1.8

(3)	um parecer fundamentado por escrito ao primeiro- secretário,	o qual, escusado seria dizer,	passava a encomenda ao segundo-secretário,	este ao terceiro,
<i>N.Car</i>	NP <sub>i</sub>	WH <sub>i</sub>	pro <sub>i</sub> V	DEM <sub>j</sub>
	58	29	41	16
TRT/Char	<b>0.046</b>	<b>0.050</b>	<b>0.049</b>	<b>0.034</b>
	0.016	0.020	0.018	0.021
MFD	<b>0.288</b>	<b>0.254</b>	<b>0.262</b>	<b>0.297</b>

	0.053	0.077	0.047	0.119		
NFix	<b>9.3</b>	<b>6.3</b>	<b>7.8</b>	<b>2.1</b>		
	2.4	3.4	2.5	1.1		
(4)	até chegar outra vez à mulher da limpeza,	que despachava sim ou não	conforme estivesse de maré.			
	NP <sub>i</sub>	WH <sub>i</sub>	pro <sub>i</sub>			
N.Car	30	25	26			
TRT/Char	<b>0.034</b>	<b>0.046</b>	<b>0.061</b>			
	0.010	0.022	0.020			
MFD	<b>0.234</b>	<b>0.259</b>	<b>0.290</b>			
	0.038	0.065	0.066			
NFix	<b>6.2</b>	<b>4.6</b>	<b>5.8</b>			
	1.9	2.2	1.9			
(5)	O homem	que queria um barco	levantou-se do degrau da porta	quando começou a abrir os ferrolhos	enrolou a manta	e pôs-se à espera
	NP <sub>i</sub>	WH <sub>i</sub>	NP <sub>i</sub>	pro <sub>i</sub>	NP <sub>i</sub>	NP <sub>i</sub>
N.Car	7	19	30	36	15	18
TRT/Char	<b>0.083</b>	<b>0.042</b>	<b>0.040</b>	<b>0.059</b>	<b>0.048</b>	<b>0.048</b>
	0.032	0.017	0.018	0.025	0.013	0.019
MFD	<b>0.259</b>	<b>0.246</b>	<b>0.265</b>	<b>0.260</b>	<b>0.367</b>	<b>0.261</b>
	0.089	0.093	0.054	0.053	0.146	0.068
NFix	<b>2.4</b>	<b>3.4</b>	<b>4.6</b>	<b>8.0</b>	<b>2.1</b>	<b>3.4</b>
	1.0	1.6	2	2.6	0.7	1.1

When we compare the mean values of RT/Char, we find that there is a progressive increase of RT from regions that include NP gaps, to those with WH gaps, culminating in the largest RT in regions with the null pronoun (Figure 1).



**Figure 1** .Mean values of RT/Char in the three linguistic variables under analysis: the three type of syntactic null categories, which occur as anaphoric elements in sentential position.

An ANOVA analysis revealed a statistical significant difference between groups ( $F(2,57)= 7,779$ ;  $p= 0,001$ ), namely between WH- ( $M=0,047$ ) and pro ( $M=0,060$ ) ( $p=0,021$ ), and NP ( $M=0,045$ ) and pro ( $M=0,060$ ) ( $p=0,003$ ).

What contribute for the observed differences in RT are the variations on the fixations in what concerns their number and duration.

## 4. GENERAL DISCUSSION

The differences between the gaps and *pro* are significant; however, the differences between the types of gaps are not markedly relevant. If we consider the computational steps involved in the grammatical derivation of the sentences presenting these syntactic objects – *pro*, variable and NP-gap – and the linguistic sources of information required for anaphoric chain resolution, we have, in fact, different conditions of processing, which have been captured by eyetracking.

The null pronoun in Subject position requires that the processor must search for an appropriate linguistic antecedent that can be found in or out of the sentence. In fact, *pro* is a pronoun, and it is free to obtain its referential content from the sentence or from the discourse (Chomsky, 1981). In the case under analysis, *pro* enters in a coreferential chain, and its antecedent is distant in structural terms: it is in other distinct clause. (Chomsky 2001, 2008). When the antecedent is found, the processing system must link the two elements and integrate them in a coreferential chain, suitable to the ongoing mental model.

We find evidences that, underlying the processing of *pro* and the two other gaps, there are different processes, which result from distinct constraints. When a WH phrase (*What, Who, When,...*) is found, it is taken as a warning to the processor to immediately search for a gap in structure building. The syntactic mechanism for parsing is committed to searching for a gap to the WH filler (assuming the *Active Filler Strategy* proposed by Frazier 1987; Frazier and Clifton 1989, 1996). When the gap is found, linking and integration take place to interpret it as a copy of the WH phrase in the left periphery of the sentence without any need of consulting the mental model.

In the case of NP gap, the linguistic processes are slightly different. In fact, the gap is not announced to the processor by a linguistic filler, and the linking strategy is guided only by the argumental structure of the verb, which is not saturated unless its external argument (the Subject) is recovered.. Besides, the structural distance between the anaphoric element and its antecedent is shorter in the case of NP copy than in the case of the WH-copy, since the former must be related with a binder in an argumental position, while the latter is bound by a wh-phrase in the periphery of the sentence.

## 5. CONCLUSIONS

In sum, anaphoric relations with coreferential *pro* are costlier than those involving gaps arising from movement. These costs can be credited to structural and discursive conditions: the referential content of *pro* is recovered from an antecedent that is distant in derivational and structural terms; its identification requires additional grammatical and discursive costs. Gaps from movement have fewer costs than *pro*, because its resolution involves only grammatical information and the structural distance between the two chain members is shorter, they are in the same syntactic domain, what does not happen with coreferential chains. In the particular case of NP copy there is a shortening of the distance between chain members because there are fewer steps involved in the derivation than in WH- chains.

To study the syntactic and semantic linguistic conditions that include anaphoric chains, we have raised the hypothesis that the higher the processing costs imposed by the structures are, the more cognitive resources will be allocated. The increase of the cognitive effort would be visible in an increase of the number or/and the durations of eye fixations. We have confirmed this prediction from the point of view of grammar, the three syntactic null categories are ranked as predicted. The covert pronoun *pro* imposes more processing costs than variables and copies, that is to say that there is an increase of costs in processing categories with more referential weight and structurally more distant from their antecedents.

In this particular case, contrary to what happens in the study of other linguistic conditions, like ambiguous words or structures, we want to measure visual reactions to an element that has no phonological form. The option by the analysis of textual regions that include these null forms is the only way of doing it. The fact that we delimited regions whose left border begins with these null forms allow us to accept that variations in its right side will be a sign of the processing of the whole structure, that is, a clause whose Subject is the target null category.

In subsequent research, using eyetracking, we will explore other measures that can highlight the results and conclusions. For example, the analysis of gaze duration (the sum of all fixations in a region) or the analysis of regressive and progressive saccadic movements to control the first and second pass,

will give us information about the reader's needs to return to some critical structural points that are crucial to explain and comprehend better the coreferential processing.

## 6. REFERENCES

- Alonso-Ovale, L., Fernandez-Solera, S., Frazier, L., & Clifton, Ch. (2002). Null vs. overt pronoun and the Topic-Focus articulation in Spanish, *Rivista di Linguistic*, 14(2), 1-19.
- Ariel, M., (1996). Referring expressions and the +/- coreference distinction. In J. Gundel, & T Fretheim (Eds.), *Referent and Referent accessibility* (pp. 13-35). Amsterdam: John Benjamins.
- Ariel, M., (2001). Accessibility theory: An overview. In T. Sanders, J. Schliperoord, & W. Spooren, (Eds.), *Text representation* (pp. 29-87). Amsterdam: John Benjamins.
- Barbosa, M. (1995). *Null subjects*. PhD dissertation. Cambridge: Mass.
- Barbosa, M., Kato, M., & Duarte, E. (2005). Null subjects in European and Brazilian Portuguese. *Journal of Portuguese Linguistics*, 4(2), 11-52.
- Batoréo, H., & Costa, A. (1998). Referência nominal na narrativa oral e escrita aos dez anos de idade. *Actas do XIII Encontro Nacional da Associação Portuguesa de Linguística*, 1, 137-149.
- Batoréo, H. (2006). *Expressão do espaço no Português Europeu. Contributo psicolinguístico para o estudo da linguagem e da cognição*. FCG/FCT, Lisboa.
- Boland, J.E. (2004). Linking eye movements to sentence comprehension in reading and listening. In M. Carreiras & Ch.Clifton, (Eds.), *The online study of Sentence comprehension*. Psychology Press, New York.
- Brito, A. (1991). Ligação, co-referência e o princípio evitar o pronome. In *Actas do Encontro de Homenagem a Óscar Lopes, Porto*, 101-121.
- Carminati, M. (2002). *The processing of Italian subject pronouns*. PhD dissertation. University of Massachusetts at Amherst. GLSA publications, Amherst.
- Chomsky, N. (1981). *Lectures on government and binding*. Foris, Dordrecht.
- Chomsky, N. (2001). Derivation by phase. In M. Kenstowicz & K. Hale (Eds.), *A life in language* (pp. 1-52). Massachusetts: The MIT Press.
- Chomsky, N. (2008). On Phases. In R. Freidin, C. Otero & M. L. Zubizarreta (Eds.), *Foundational Issues in Linguistic Theory* (pp. 133-166). Massachusetts: The MIT Press.
- Colaço, M. (2005). *Configurações de coordenação aditiva: tipologia, concordância e extracção*. PhD dissertation. FLUL, Lisboa.
- Costa, A. (2010). Referential chain's processing. In *Papers in Psycholinguistics – Proceedings of the First International Psycholinguistics Congress of ANPOLL* (A França and M. Maia, Orgs). Rio de Janeiro: Ed. Imprinta.
- Costa, A. (2003). *Processamento de frases em Português Europeu*. PhD thesis, FLUL. [Fundação Calouste Gulbenkian, Lisboa, 2005]
- Costa, A., & Luegi P. (2009). Complexidade linguística e processamento referencial. In *Anais do VI Congresso Internacional da ABRALIN* (D. da Hora, Org.), CD-ROM, João Pessoa.
- Costa, A., Faria, I., & Kail, M. (2004). Semantic and Syntactic Cues' Interaction on Pronoun Resolution in European Portuguese. *Proceedings of the 5th Discourse Anaphora Resolution Colloquium*, Lisboa, 45-50.
- Costa, A., Faria, I., & Matos, G. (1998). Competitive information sources in referential ambiguity resolution. *Proceedings of 5th International Congress of the ISAPL*, Porto, 173-187.
- Costa, A., Matos, G., & Luegi, P. (2010). Processamento de relações anafóricas com sujeitos omitidos em Português Europeu. In *Textos seleccionados do XXV Encontro Nacional da Associação*



- Portuguesa de Linguística* (pp.351-364). Lisboa: Associação Portuguesa de Linguística.
- Featherston, S. (2001). *Empty categories in sentence processing*. Amsterdam: John Benjamins Publishing Company.
- Fodor, J. (1993). Processing empty categories: a question of visibility. In G. Altmann & R. Shillcock, (Eds.), *Cognitive models of speech processing: the Second Sperlonga Meeting* (pp. 351-400). Hillsdale: Lawrence Erlbaum Associates
- Frazier, L.(1987). Syntactic processing: Evidence from Dutch. *Natural Language and Linguistic Theory*, 5,519-559.
- Frazier, L., & Clifton, Ch. (1996). *Construal*. Cambridge: A Bradford Book.
- Hickmann , M.(1995). Discourse organization and the development of reference to person, space and time. In P. Fletcher & M. Garman (Eds.), *The Handbook of Child Language*. Cambridge: CUP.
- Hyon, J ., Lorch R., & Rinck, M. (2003). Eye movement measures to study global text processing. In J. Hyon, R. Radach & H. Deubel,(Eds.), *The Mind's Eye. Cognitive and applied aspects of eye movement research*. Amsterdam: Elsevier Science BV.
- Just, M., & Carpenter, P. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, 87, 329-354.
- Mateus, M., Brito, A., Duarte, I., Faria, I., Frota, S., Matos, G., Oliveira, F ., Vigário M., & Villalva A. (2003). *Gramática da Língua Portuguesa*, Caminho, Lisboa.
- Matos, G., & Duarte I. (1984). Clíticos e sujeito nulo no português: contribuições para uma teoria de pro. *Boletim de Filologia*, 29(2), 479-538.
- Matos, G. (2000). Across-the-board clitic placement in Romance languages. *Probus*, 12(2), 229-259.
- Rayner, K., & Pollatsek, A. (2006). Eye-movement control in reading. In M. Traxler and M. A. Gernsbacher, (Eds), *Handbook of Psycholinguistics*. Oxford: Elsevier.